Dissecting the Complex Diseases

Geneticists face a tough challenge in the 1990s as they begin efforts to tease out the genetic components of multi-cause diseases such as cancer and schizophrenia

SCIENCE IN THE '90s

Seventh in a series

DURING THE 1980S, molecular geneticists took a major step toward understanding the simple hereditary diseases, those caused by single-gene mutations. Counted among their successes, for example, were the isolation of the long-sought genes that cause Duchenne muscular dystrophy and cystic fibrosis.

But remarkable as the achievements are, they will have only a modest impact on public health, for the simple genetic diseases are relatively rare, numbering their victims in, at most, the tens of thousands. So as the 1990s begin, molecular geneticists are turning to the far more daunting task of unraveling the genetic components of the much more complex conditions that cause the vast majority of death and illness in the developed world. "Solving the single-gene diseases is straightforward—hard work—but straightforward," says geneticist Eric Lander of the Whitehead Institute in Cambridge, Massachusetts, and one of the moving forces in the new effort. "The frontier is diseases of more complex inheritance."

Those diseases include heart disease and cancer—the two leading killers—as well as high blood pressure, obesity, and common autoimmune disorders, such as diabetes and multiple sclerosis. And there are also mental diseases, such as Alzheimer's, manic-depression, and schizophrenia. There is good evidence from both animal studies and human observation that genetic makeup influences susceptibility to these diseases. In fact, susceptibilities to many of the complex diseases seem to be influenced by perhaps three to five genes acting together.

Genetic susceptibility is apparently not the only determinant of whether an individual actually gets sick, however. One indication that it's not comes from studies of identical twins, who are born with the same genes. The studies show that when one twin gets a disease, such as diabetes or schizophrenia, the other one gets it no more than 20 to 50% of the time. Environmental influences, such as diet, smoking, chemical exposures, viral infections, and the all too common "unknown" must also play a role.

All that complexity will make the job of teasing out the genetic components of the complex diseases an uphill slog. It certainly won't be completed in this decade. But the hope is that the effort will ultimately pay off by giving researchers a better grasp of how these diseases develop. And on a practical level, a more thorough understanding of the diseases might lead to better therapies and perhaps to better prevention.

One indication of the intense current interest in the complex diseases was a meeting on the "Genetics and Molecular Biology of Complex Diseases," held earlier this month at the Banbury Center at Cold Spring Harbor Laboratory on Long Island. Banbury meetings usually focus on a specific disease or area of investigation, says Lander, who cochaired the confab with Aravinda Chakravarti of the University of Pittsburgh and Banbury Center director Jan Witkowski. But the point of this one was to get an exchange going among people working on different disorders—to attack in common the immense technical problems posed by the complex diseases.

One problem they all face is that the "experimental organisms" are human beings. "There is a major difference between doing genetics like this in humans and in other animals or plants," Lander says. "You can't do crosses. You can't manipulate the environment."

What the researchers often try to do instead is use linkage analysis, a genetic technique that was given a big boost by the advent during the past decade of the chromosomal markers called restriction fragment length polymorphisms, or RFLPs ("riflips"). RFLPS are nothing more than short DNA sequences that vary from one person to the next and can be mapped to specific chromosomal locations.

The idea is to find RFLPs that show a lot of variation and then use them in studies of families who have genetic diseases to see if family members who get the disease consistently carry a particular variant. If they do, the researcher can conclude that the disease gene and the RFLP are "linked": they are inherited together and therefore must be located very near one another on the same chromosome. The RFLP marker is then used as a stepping-stone to find the gene itself. The researchers who found the cystic fibrosis and muscular dystrophy genes used this approach, which does not require any knowledge about the nature of the target gene or its chromosomal location.

But applying RFLP linkage analysis to the complex diseases is going to be much more difficult. Here researchers are looking for genes that may contribute only 10 or 20% of an individual's susceptibility, rather than 100%, as in the case of the single-gene disorders. Such genes may be hard to detect by randomly scanning the genome for an association between a disease and a RFLP marker. In fact, recent fiascoes with linkage analysis of two complex disorders—manic-depression and schizophrenia—illustrate some of the hazards.

In November 1987, Janice Egeland of the University of Miami and her colleagues reported that they had linked cases of manic depression in a large Amish family to RFLP

COMMON DISEASES WITH COMPLEX CAUSES

Disease	People affected in the United States	Candidate genes (*confirmed)
CANCER	6 m (1 m new cases expected in 1990)	Oncogenes, tumor suppressor genes. *Retinoblastoma, Wilms' tumor genes
CORONARY HEART DISEASE	5 m (1.5 m heart attacks expected in 1990)	Genes involved in cholesterol and lipid metabolism, blood clotting. *LDL receptor gene
DIABETES	0.5 m	*Histocompatibility protein gene. Genes for T cell receptor, other immune molecules
HIGH BLOOD PRESSURE	60 m	Renin gene
MANIC-DEPRESSION	1 m	?
SCHIZOPHRENIA	1.8 m	?

markers on chromosome 11, raising hopes that a gene for the disorder would be found soon. But those hopes were dashed at the end of last year, when the linkage report had to be retracted after further analysis failed to confirm it (also see *Science*, 17 November 1989, p. 886).

Schizophrenia workers have gone through a similar boom-and-bust cycle. At the end of 1988 Hugh Gurling of University College and Middlesex School of Medicine in London and his colleagues reported linkage between schizophrenia and markers on chromosome 5. Accompanying the Gurling report, however, was another in which Kenneth Kidd of Yale University School of Medicine and his colleagues said that they failed to find such a linkage.

Now, the two groups might have come up with different results because they did not use the same families for their linkage analyses, leaving open the possibility that a gene on chromosome 5 contributes to the schizophrenia of the family studied by the Gurling group, but not to that of the family studied by Kidd and his colleagues. Indeed, researchers expect to find this type of genetic heterogeneity in the complex diseases, although heterogeneity cannot be definitively claimed unless linkage to a second gene is actually found.

The chromosome 5 linkage still has not been confirmed, however, and Kidd is not alone when he says, "My personal opinion is that [Gurling's] finding is a false positive. Not that he did anything wrong, but that it was just a statistical fluke." Kidd notes that he was a member of the original manic-depression study and can't "throw rocks."

Why are researchers finding the mental diseases so difficult to study by linkage analysis? Two participants in the Banbury meeting, research psychiatrists Miron Baron of New York State Psychiatric Institute in New York City and Irving Gottesman of the University of Virginia School of Medicine in Charlottesville, pointed to one major complicating factor. Manic-depression and schizophrenia are extremely hard to diagnose accurately.

And large numbers of patients do not have to be misdiagnosed to confound linkage analysis. In the manic-depression study, for example, diagnosis of the disease in just two additional members of the Amish family went a long way toward wiping out the original finding.

After hearing about the difficulties with the work on manic-depression and schizophrenia, Mary-Claire King of the University of California, Berkeley, whose research focuses on breast cancer, said, "There's no question that this is as hard as it gets."

But the embarrassing fiascoes with manic-depression and schizophrenia do not mean that linkage analysis of complex diseases is hopeless. Molecular geneticists expect that they will succeed with other diseases in which accurate diagnosis is not the problem that it is for the mental diseases. They are especially optimistic about conditions for which they have good biological indicators that can be used in looking for linkage.

One such condition cited by several geneticists is coronary artery disease, the cause of heart attacks. About 10 years ago, Joseph Goldstein and Michael Brown of the University of Texas Southwestern Medical Center in Dallas (who shared the 1985 Nobel Prize for Medicine for their work) identified a gene mutation that results in extremely high concentrations of LDL (low-density lipoprotein) cholesterol, a type of cholesterol that is associated with a high risk of having a heart attack. In fact, people who get two copies of the mutation often die of

From the Tomato to the Mouse

Learning how to grow better tomatoes wouldn't seem to have much to do with current efforts to understand the genetics of complex diseases such as cancer, diabetes, and schizophrenia. But it does, according to geneticist Eric Lander of the Whitehead Institute for Biomedical Research. The traits that interest tomato fanciers—size and solid content, for example—are influenced both by groups of genes acting in concert and by environmental factors, as is development of the complex diseases. Therefore, Lander proposes, a new method for mapping the chromosomal locations of the genes that specify complex traits—a technique that has so far been used only in the tomato—might help to dissect out the genes that make people susceptible or resistant to complex diseases.

To find the genes that influence particular characteristics, geneticists do breeding studies in which the inheritance of the trait under study is correlated with "markers" having known chromosomal locations. If a marker and a trait are consistently inherited together, they must be located close together in the genome. In the past, however, finding markers for mapping the locations of the genes determining quantitative traits was strictly a matter of luck. But "now we no longer need to be lucky to map quantitative trait loci," says Lander. His optimism—and the new method—rest on the development of the chromosomal markers called RFLPs. "With RFLPs, Lander adds, "we can mark the entire genome."

RFLP linkage analysis has already been widely used for finding single genes—the cystic fibrosis gene, for example. What Lander and David Botstein of the Stanford University School of Medicine have done is to devise analytical and mathematical methods for extending the approach to map the multiple genes that collectively specify complex traits. Lander's group, working in collaboration with that of Steven Tanksley of Cornell University in Ithaca, then tested the method in the tomato. The researchers identified at least six chromosomal locations encoding genes for fruit mass, four encoding genes for soluble solids concentration, and five encoding genes for acidity. The collaborators are now trying to zero in on individual genes at those sites.

Lander is the first to concede that the method for mapping quantitative trait loci is not applicable to humans because it depends on doing breeding experiments. But it should, he says, work with rats and mice, which have a genetic complexity similar to that of the tomato. Lander and hypertension expert Victor Dzau of Harvard Medical School are going to try to use the method to identify the locations encoding the genes that regulate blood pressure in the rat. Once they have candidate genes, they can go on to see whether the same genes affect blood pressure in human beings, too. ■ **J. M.**

ADDITIONAL READING

E. S. Lander and D. Botstein, "Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps," *Genetics* **121**, 185 (1989).

A. H. Patterson *et al.*, "Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms," *Nature* **335**, 721 (1988).

heart attacks in their twenties.

So the LDL cholesterol concentration provides an easily measured trait that can be used as an indicator to search for genetic factors that make people more or less susceptible to heart attacks. Recently, for example, Helen Hobbs, who works in Goldstein's lab in Texas, found evidence in a linkage analysis study for a gene that acts to suppress blood LDL concentrations and might possibly protect against heart attacks, although that remains to be established.

The Texas workers are currently attempting to find the LDL-suppressing gene. Their search has not yet been successful, but geneticists are confident that it will be.

Linkage analysis is performed with families, as Hobbs' study and the ill-fated schizophrenia and manic-depression genes were. But there is another genetic technique called association analysis that is also being put to use in the effort to find disease susceptibility genes. Association analysis looks for connections between genes and hereditary traits in populations. And because population members are much more genetically diverse than family members, the researchers have to narrow down their search by focusing on candidate genes for which there is biological reason to think they might be involved in a disease.

Insulin-dependent diabetes provides a case in point. This form of diabetes is an autoimmune disease in which the immune system attacks and destroys the beta cells of the pancreas, which secrete the body's insulin. Immunologists have known for many years that people who get diabetes carry certain variants of the histocompatibility proteins. Because the histocompatibility proteins are critical regulators of immune cell reactions, the assumption was that the presence of the particular variants was somehow related to the development of the aberrant immune response.

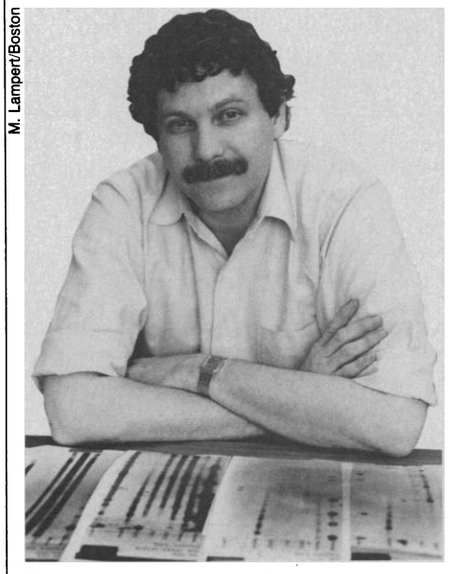
That supposition was borne out by further work. About 3 years ago, for example, Hugh McDevitt's group at Stanford University School of Medicine narrowed diabetes susceptibility down to a single amino acid in one of the histocompatibility proteins. The presence of aspartic acid at position 57 in the protein tends to protect against diabetes, the researchers found, whereas having another amino acid there increases a person's risk of coming down with the disease.

The original associations between histocompatibility proteins and the autoimmune diseases were made without the aid of RFLP markers because the histocompatibility proteins naturally vary from one person to another. But the availability of RFLPs has opened up the possibility of doing association analysis for other candidate genes that lack such natural variability. To do this, the researcher looks for a RFLP marker located very near the gene under investigation, perhaps in one of the noncoding regions contained within most of the genes of higher organisms, and then determines whether the RFLP is associated with a disease.

Steve Humphries of Charing Cross Sunley Research Center in London described one such analysis at the Banbury meeting. High levels of the clotting factor fibrinogen seem to predispose people who have already had one heart attack to having another. Humphries has now found a RFLP in the regulatory region of the fibrinogen gene that goes along with having high concentrations of the clotting factor. "RFLP studies are a useful first tool," Humphries concludes, "in an initial screen for functionally important sequence differences that may be related to disease."

Comparable studies might also be done with other complex diseases. Cancer, for example, has a large number of candidate genes. These include genes encoding enzymes that break down chemicals and drugs, plus the 50 or so "proto-oncogenes," genes that can bring about the cancerous transformation of cells if they undergo certain mutations, and a growing number of tumor suppressor genes, the loss of which may also contribute to cancer development.

Most of the gene changes that have been identified in cancerous tumors are thought to arise after birth, presumably as a result of lifetime exposures to environmental carcinogens, but there are exceptions to this. Hereditary defects in some suppressor genes predispose to certain childhood cancers, including retinoblastoma and Wilm's tumor. The quest now is to see if inherited forms of oncogenes or suppressor genes make people susceptible to other cancers as well.



M. Lampert/Boston

Tracer of lost genes. *Geneticist Eric Lander tracks complex traits in tomatoes and humans.*

What does the coming decade hold for the investigation of the complex diseases? It seems clear that this work is still in its infancy—although it is gathering steam—and that some key technical advances will be needed for it to mature. One thing that will help is improved methods for doing genetic analysis of complex traits. Lander and David Botstein of Stanford University School of Medicine may have developed one such method, although it is not applicable to humans (also see box on p. 1541).

Another thing researchers would dearly like to have is a detailed marker map of the human genome to help them in their quest. That would make it easier, Lander says, to follow the inheritance of a chromosome in a population and improve the correlations between the markers and the trait under study. A genetic map is supposed to be one of the priorities of the Human Genome Project, but the program has taken a lot of criticism recently for letting the map slip (*Science*, 19 January, p. 281).

The researchers also want to identify families in which a particular disease is passed from generation to generation. The National Institute of Mental Health is helping by setting up a multimillion-dollar 5-year program for identifying families with schizophrenia, manic-depression, and Alzheimer's disease. An early goal of the NIMH program is to devise standardized diagnostic criteria for these conditions to help reduce the scientific complications that are introduced by inaccurate diagnoses.

Researchers interested in unraveling the genetic components of the complex diseases clearly have their work cut out for them in the coming decade—and no doubt beyond. The task will be much more difficult than solving the single-gene diseases, which are hard enough. Is it worth the trouble? Psychiatrist Kidd was speaking of the mental illnesses, but his words apply to the other complex diseases as well: "These conditions are so important both as public health problems and individual health tragedies that they warrant major efforts to try to understand them," he says. ■ JEAN MARX